DOCUMENT RESUME

ED 214 137 CS 006 556

AUTHOR Daniel, Mark
TITLE Vocabulary-Test Errors and Word Learning. Technical Report 1981-8.
INSTITUTION Johnson O'Connor Research Foundation, Boston, MA.
PUB DATE Dec 81
NOTE 25p.

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Adults; *Error Analysis (Language); Evaluation Criteria; *Language Acquisition; *Learning Theories; Reading Instruction; *Reading Research; *Validity; *Vocabulary Development; Vocabulary Skills

ABSTRACT

Vocabulary test errors made by people with varying levels of vocabulary knowledge might give information about the stages involved in learning a word's meaning. A study used such data to evaluate Johnson O'Connor's proposal that word learning involves four stages and that each stage is characterized by a type of confusion (a mislead). In the first or "look alike" stage, people tend to confuse a word with other words similar in sound or appearance, while in the second or "context" stage, people tend to confuse a word with others in the same setting. In the third stage, one might have a greater knowledge of word meaning, but confuse it with its exact opposite, or antonym. The last or close stage would involve fine shades of meanings of a word. Subjects were 326 adults who took a 150-item vocabulary test that included examples of each of the four O'Connor categories. The results showed only a small degree of relationship between mislead categories and vocabulary level, and did not support the O'Connor proposal regarding the way misleads are ordered. (FL)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED214137

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

This document has been reproduced as received from the person or organization originating it.

Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

VOCABULARY-TEST ERRORS AND WORD LEARNING

by

Mark Daniel

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

R. Bowker

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

JOHNSON O'CONNOR RESEARCH FOUNDATION

HUMAN ENGINEERING LABORATORY

Technical Report 1981-8

December, 1981

CS006 556

Vocabulary-Test Errors and Word Learning

Mark Daniel

SUMMARY: Errors on vocabulary tests made by people with varying levels of vocabulary knowledge might give information on the stages involved in learning a word's meaning. The present study uses such data to evaluate Johnson O'Connor's proposal that word learning involves four stages: associating the new word with others which sound or look like it; associating the word with its appropriate context; knowing the general meaning of the word but confusing it with its exact opposite; and knowing the precise meaning of the word. An analysis of test items having misleads corresponding to the four stages shows that, contrary to O'Connor's hypothesis, confusion of a word with its opposite tends to occur at a very early stage of word learning. There is only a very weak relationship between mislead type and vocabulary level; this is probably due in part to the effects of item idiosyncracies on answer choices.

Many psychologists interested in language learning have shown an increased interest in recent years in the development of semantic knowledge. Following a period in which syntax and grammar were the major objects of attention, the past five or ten years have produced an increasing number of studies of how words and their associated concepts are acquired, stored and recalled. One of the most productive theories in this domain has been the "network" theory of concepts (Wickelgren, 1981), in which a concept is seen as a "node," or locus, of connections with other concepts. In this theory, the process of acquiring a new concept is the process of strengthening the connections or associations among a set of existing concepts.

Some work has investigated the steps involved in the formation of verbal concepts, particularly among young children. For example, Donaldson and Wales (1970) observed that the positive member of a pair of opposite words (such as "more/less") was learned before the negative member. Clark (1973) reported similar results with relational pairs of words such as long/short, wide/narrow and high/low, in which the "positive" member of the pair is the one which can be used to refer to the dimension (e.g. "How long is it?"). From these studies is is inferred that the positive concept is formed first, and the negative concept is formed later when connections are established with the concept "not." On the other hand, Kurcaj and Maratsos (1975) showed that the words "front" and "back" are learned simultaneously, even though (because?) "back" is defined as "the other side from the front." Perhaps the spatial nature of these terms contributes to their being learned simultaneously rather than sequentially as with such non-spatial relational pairs as more/less or big/little.

Another question in the field of word-learning is whether the early stages of knowing a word are primarily characterized by overgeneralization or under-

generalization. That is, does the set of referents of a word expand or contract as the child's understanding of the word becomes more precise? Early child-diary studies suggested that overgeneralization was typical in the first stages of word learning, but Anglin (1977) showed how these diaries are biased in favor of finding overgeneralization. Saltz, Soller and Siegel (1972) argued the opposite position, that the progression in children's concept learning is from fragmented, overlapping subconcepts to more integrated concepts. More recently, a number of investigators have concluded that the learning of a verbal concept may involve increasing generalization or increasing specificity, depending on the concept. Niemark (1974), for example, found that the concept "food" became more broadly defined with age whereas the concept "clothing" became more narrowly defined. Anglin (1977) argues that the functional aspects of words (that is, their usefulness to the child) are the major determinants of the initial level of generalization, and Rosch (1977) suggests that the earliest words learned will usually be of intermediate generality.

O'Connor (1940), who was primarily concerned with the measurement of vocabulary knowledge among teenagers and adults, suggested on the basis of informal item analysis that there is a generally applicable sequence of stages in word learning, each stage being characterized by a particular type of confusion. At an early stage, people tend to confuse the word with other words that are similar in sound or appearance. As knowledge increases, the word tends to be confused with other words used in the same context or setting. As illustration, O'Connor gives the following: "Children see a box and hear it called a GIFT, or receive a gift and hear it called a BOX, and mark one as the meaning of the other" (p. 126). In a later stage of learning, one has a greater knowledge of the word's meaning but is ilkely to confuse the word with its exact opposite; for example, one may

know that "port" and "starboard" refer to left and right, but be uncertain about which is which. The last stage would be a highly refined understanding of the word, with confusions being likely to involve fine shades of meaning (for example, distinguishing between "forecast" and "analysis" as the better synonym of "prognosis").

This description of the process of word learning is consistent with the network theory of verbal concepts. As a concept becomes elaborated, it develops strong connections with an increasing number of other concepts. Initially the concept denoted by the new word has only a vague meaning because it has only a small number of strong connections with other concepts. The connections may often have to do with the context in which the new word has been seen or heard. As more is learned about the word, more connections are strengthened. Two words which are exact opposites of each other might conceivably have networks which are the same except that a negative modifier is involved in one of them; thus, confusions between the opposites might be expected during the more advanced stages of learning a word. Learning the subtleties of a word's meaning is a process that may continue for a long time, so we might expect a person in the late stages of learning a word to confuse it with other words whose meaning networks differ only slightly from its own.

Perhaps the most interesting aspect of O'Connor's hypothesis is the location of the antonym in the sequence. It recalls some of the research referred to earlier in suggesting that the concept of negation is incorporated relatively late in the formation of a new concept. It suggests that a word and its exact opposite are represented in a very similar or interconnected way in conceptual memory. This is certainly a questionable proposal. The connection with the concept "not" might occur at any one of a number of levels of aggregation; for example, one of the relatively basic concepts participating in the node that is

the new concept may itself carry the negative value.

The present research is designed, first, to evaluate empirically the O'Connor hypothesis by examining the types of mistakes people make on a multiple-choice vocabulary test. Each of the items being examined has misleads (or distracters) representing three or four of the following categories:

1) "sound- or look-like" mislead: synonym of a word sounding or looking like the test word;

2) "context" mislead: a word used in the same context as the test word, but not being similar in meaning;

3) antonym; and

4) "close" mislead: a word which is similar in meaning to the test word but not as exact a synonym as the correct answer.

We will be looking for evidence that the attractiveness of these mislead categories changes as examinees' understanding of the test word increases, in the specific order specified by O'Connor.

A second objective of this study is to find <u>any</u> ordering of mislead categories which is related to word knowledge. The O'Connor hypothesis is only one of many possible orderings; whether it is supported or rejected by the data, we will want to know if there is some other ordering that might be superior.

The question of mislead ordering is of interest from a psychometric point of view. If there is a generalizable ordering of mislead categories which is related to examinees' knowledge of the test words, then mislead choices would convey information beyond that provided by whether or not the examinee chose the correct answer. Incorporating this information into the total score ought to increase reliability. The usual method for utilizing mislead information is called "item-option weighting", and such weighted scoring systems have been found to increase reliability substantially (Claudy, 1978; Downey, 1979). However, item-option weighting usually uses empirically-derived weights, and so is

costly and time-consuming to implement. If there were a known ordering of mislead categories for vocabulary tests, the item weighting could be simplified to the point that the scoring of misleads might be practical.

METHOD

Materials

This study uses a five-choice, 150-item, untimed test of general English vocabulary developed by the Johnson O'Connor Research Foundation, Inc. (Work-sample 641 form EC). This is one of the hardest in a graded series of narrow-range tests; the complete series is intended for use from grade 7 through the high-vocabulary adult level. The printed instructions preceding the test are as follows:

> Each test word is printed in italic in a phrase. From the five choices to the right of the phrase, select the one which comes nearest in meaning to the meaning of the test word. Underline the one you select. The following example is correctly marked.
>
> Drive the *automobile* locomotive herd robot glider car

The 150 items of the test are arranged in increasing order of difficulty, based on the proportion of examinees in previous samples answering each item correctly.

Subjects

The examinees are 326 clients of the Johnson O'Connor Research Foundation. These individuals, most of whom are college students or adults, paid a fee to take a wide-ranging battery of aptitude and knowledge tests, of which the vocabulary test was a part. Their reason for taking the test battery was to gain information that might help them in making career and educational decisions. All were tested in Boston, Massachusetts or Fort Worth, Texas. The age range is from 16 to 55 (mean: 25.8, standard deviation: 9.2); 59% are male. Their average vocabulary-test score is predictive of a score of 609 on the verbal

portion of the Scholastic Aptitude Test (using a linear regression equation based on 100 high-school juniors and seniors, for whom the tests correlate .85; Bowker, 1975).

Procedures

A short vocabulary placement test was used to select those examinees who would take the particular narrow-range test used in this study. All tests were administered with instructions to answer every item, even if this required guessing. This instruction to guess is of value in this study because we are interested in the mislead choices of people who have little or no knowledge of the meaning of the test word.

Selection of Items for Analysis

When the 641 series of tests was constructed, an effort was made to include an example of each of the four mislead categories in each item. However, many of these misleads were replaced during revisions based on item analyses, usually because they did not attract many responses. The test developers did not document the categories that they intended each mislead to represent. Therefore, it was necessary for the present author to examine each of the 150 items to see if at least three of the four mislead categories were included, using the following criteria for categorization:

a) "Sound- or look-like" mislead: a synonym of a word that looks or sounds like the test word. Such misleads are easy to recognize if they are present.

b) "Context" mislead: a word whose use would be appropriate in the same situation as the test word, even though having a different meaning. There is some overlap of meaning between the "context" mislead and the test word, because the "context" mislead will be capable of referring to, or being done by or to, the same thing, person, or action as the test word. If, for example, the test word is an adjective usually applied to people, then the "context" mislead will also be an adjective applied to people. However, the "context" mislead is not nearly as similar in meaning to the test word as is the "close" mislead.

c) Antonym: most adjectives and adverbs have a direct opposite. Nouns which have a common adjectival form--such as "geniality" and "amplitude"--also have straightforward opposites. With concrete nouns, the antonym is harder to define: several different antonyms might be possible for the same test word, each opposing a different aspect of the test word's meaning. For example, "volunteer" and "craftsman" are both antonyms of "hireling," but for different reasons. Verbs, like nouns, may have a clear opposite, especially if they are frequently used in an adjectival form (such as "vexed") or if they have a negative prefix ("misapprehend"), but often they lack a direct antonym, and several different words might oppose some portion of the test word's meaning.

d) "Close" mislead: a word that has much in common with the test word, and that might even suffice as a correct answer were there not another option which is a more exact synonym. A close mislead may designate a general concept of which the test word is a subconcept: for example, the close mislead for "proselytize" is "talk", and the close mislead for "geniality" is "mood". The close mislead may be a concept at the same level of generalization as the test word and closely associated with it, but not a precise synonym; examples would be "ponderous/soaked", "enigmatical/profound", "gaunt/stooped" and "ennui/weakness".

Only 26 of the 150 items were found to contain at least three of the four mislead categories, and these items are listed in Table 1. The fact that so many of the items no longer contain even three of the four mislead categories illustrates the extent to which the effectiveness of a mislead depends on specific item characteristics; it may also suggest that some types of misleads do not represent common types of misunderstandings of many words. The 26 items that meet the criteria are widely dispersed through the 150-item test, and so are as varied in level of difficulty as the narrow-range nature of the test will permit.

Analysis

The first analysis was a direct numerical test of the O'Connor hypothesis. The hypothesis predicts that, among people who make an error on an item, those who have a better knowledge of the test word will tend to choose certain types of misleads. It is reasonable to assume that those who have a higher total

Table 1
Test Items, Showing Categorization of Answer Choices

Test word in phrase	CATEGORIES OF ANSWER CHOICES				
	Synonym	Close Mislead	Antonym	Context	Sound- or look-like
1. DUBITATIVE manner	unsure	groping	confident	distin-guished	excitable
2. Several HIRELINGS	unskilled-workers	peasants	volunteers, craftsmen	--	blackbirds
3. It VEXED him	annoyed	confused	comforted	scorned	--
4. He MISAPPREHENDED	erred	misled	grasped-correctly	guided	punished
5. The GAUNT man	thin	stooped	big	aged	--
6. The PONDEROUS material	heavy	soaked	airy	--	marvelous, debatable
7. It DETERRED him	hindered	inter-rupted	spurred	bothered	resched-uled
8. To be UNMOLLIFIED	unpacified	unfitted	easily-led	changed	liquefied
9. Her UNEQUIVOCAL stand	definite	supreme	uncertain	unimportant	--
10. DESPOILED cities	plundered	littered	prosperous	dying	over-indulged
11. His CORPULENT uncle	stout	muscular	trim	--	dead, rich
12. HEINOUS environ-ment	wicked	menacing	good	oppresive	--
13. Received his STIPEND	wages	gift	penalty	drink	writing-instrument
14. GENIALITY of the group	cheer-fulness	mood	bad-temper	mixing	ancestors
15. Their CORPORATE life	united	business	individual	middle-class	abundant
16. The VERNACULAR style	native	modern	foreign	wordy	--
17. Always with COMMISERATION	pity	expression	insensi-tiveness	gossip	greed

Table 1 (continued)

Test word in phrase	CATEGORIES OF ANSWER CHOICES				
	Synonym	Close Mislead	Antonym	Context	Sound- or look-like
18. The ENIGMATICAL disclosure	puzzling	profound	well-expressed	calculated	practical
19. Looking at the CORNICE	top-molding	joint	base	wall	angle
20. The FEINT	pretense	side-step	sincerity	charges	stain
21. Affected by ENNUI	boredom	weakness	zest	envy	poverty
22. PRETERNATURALLY active	abnormally	mysteri-ously	typically	disturb-ingly	early
23. INDEFEASIBLE right	permanent	meaning-ful	doubtful, wavering	--	unspoken
24. PRIVILY refused	secretly	consider-ately	openly	formally	--
25. Youthful PROSELYTIZING	converting	talking	thwarting	warning	--
26. Such AMPLITUDE	breadth	height	narrowness	loud-noise	inborn-gift

Note: If two different answer choices in an item belong to the same category, they are shown in the appropriate column separated by a comma.

score on the 150-item vocabulary test will, on the average, have a better knowledge of any individual word on the test. Therefore, if ascending numerical values are assigned to the mislead categories in the hypothesized order--that is, one point for the "sound- or look-like" mislead, two points for the "context" mislead, three for the antonym and four for the "close" mislead--then among people who miss an item there should be a positive correlation between total vocabulary score and mislead score. The O'Connor hypothesis would be supported if the mean of the 26 item correlations were substantially greater than zero--at least .10 (a "small" effect in Cohen's (1977) terminology). Of course, even a substantial positive correlation will not imply that all four of the mislead categories are ordered correctly in the O'Connor hypothesis; the ordering might be only partially correct and still produce a positive mean correlation.

The second analysis investigated whether the four mislead categories can be ordered in terms of their relationship to total vocabulary level. That is, is there a tendency for people choosing one type of mislead to have higher vocabulary scores than people choosing another? For each item, this question could be addressed using mean scores directly. However, comparison across items would be hindered by the use of mean scores. Item difficulty would affect the level of the means (difficult items would have higher means than easy items), and this inter-item variation would be error variance which would reduce the power of statistical analysis. A better procedure is to compute, for each item, the point-biserial correlation of each mislead category with total vocabulary score, using the sample of people who made an error on the item. The average level of these correlations is unaffected by item difficulty. The mislead category having the highest mean correlation across items is the one that tends to attract the higher-vocabulary examinees. These correlations were computed

for all 26 items, and a one-way analysis of variance was used to see whether there is significant variation among the mean correlations for the four mislead categories. The ANOVA was also applied separately to the nine nouns and fourteen modifiers (adjectives and adverbs) among the 26 test items.

RESULTS

First Analysis

In the first analysis, the mislead categories were assigned numerical values (1 through 4) and, for each item, mislead scores were correlated with total vocabulary scores, using those examinees who failed to answer the item correctly. The product-moment correlations of raw scores have a negative bias because the variables being correlated are skewed in opposite directions. The mislead-score distributions tend to be negatively skewed, because the mislead categories with higher point values tend to attract more responses. As shown in Appendix I, the "close" mislead attracts a median of 44% of the incorrect responses, whereas the "sound- or look-like" mislead attracts only 16%. The vocabulary-total distribution tends to be positively skewed for a sample of people who have made an error on an item. In order to reduce this bias, the Spearman rank-order correlation method has been used. These correlations tend to be slightly higher than the product-moment values computed for the same data in this study. The rank-order correlations are shown in Table 2.

The 26 rank-order correlations range from -.224 to .317, with a mean of .066 and a standard deviation (Fisher's z) of .133. The 95% confidence interval for the mean is (.015, .115). Thus, there is a significant tendency for the correlations to be greater than zero, but the average correlation is very small, suggesting that the O'Connor ordering of mislead categories is only slightly related to vocabulary level. Because this mean correlation is smaller than

Table 2
Correlation of Mislead Score with Total Vocabulary Score

Item	*N*	r_s	Item	*N*	r_s
1	83	.060	14	64	.153
2	42	.114	15	163	.068
3	103	.273	16	144	.023
4	55	.005	17	93	-.121
5	56	.216	18	95	-.039
6	85	.116	19	100	-.077
7	79	.146	20	98	.304
8	84	.010	21	133	.101
9	67	-.002	22	137	-.068
10	104	.245	23	164	.030
11	114	.006	24	100	-.224
12	109	-.045	25	166	.057
13	83	.317	26	211	.004

even the minimum value for a "small" effect in Cohen's system, the data will not be considered to have supported the O'Connor hypothesis.

It may also be of interest that the variance of the correlation distribution is larger than would be expected if the correlations differed from one another only because of sampling error. The square root of the mean expected variance of Fisher's *z* for the sample sizes in Table 2 is .107; the observed standard deviation, .133, is significantly larger ($F(25,\infty) = 1.55$, $p < .05$). This means that the differences among the correlation values for different items are not caused solely by sampling error, but reflect real differences in how strongly the mislead categories are related to vocabulary level. These differences may be caused by inter-item variation in the attractiveness of mislead categories; yet it is also possible that the mislead ordering is related to vocabulary level for only a subset of the items being studied.

Second Analysis

The purpose of the second analysis was to see whether some types of misleads tended to attract higher-vocabulary examinees than did other types.

This was accomplished by computing for each item the point-biserial correlation between the mislead and total vocabulary scores, for the sample of examinees who did not answer the item correctly. The results are presented in Table 3.

Table 3
Correlations of Mislead Types with Total Scores

Mislead Type	No. Items	Corr. with Total(z trans.) Mean	SD
Close	25	.064	.142
Antonym	23	-.055	.148
Context	20	-.008	.152
Sound-like	19	-.029	.137

Note: Figures are Fisher's z values (which, below .11, are identical to correlation values to three decimal places).

Analysis of variance was applied to these data to see whether the observed variation in means is statistically significant; results are given in Table 4.

Table 4
ANOVA Table for Data in Table 3

Source	SS	*df*	MS	*F*
Between	.1870	3	.0623	2.82
Within	1.8313	83	.0221	
Total	2.0183	86		

The value of *F* is significant at the .05 level. Therefore, we can be fairly confident that there is some difference among the four mislead categories in the average vocabulary levels of the examinees they attract. Two things should be noted about this result. First, the order of mislead categories is not that which was predicted by O'Connor's hypothesis. Second, the magnitude of the relationship between mislead category and correlation with total vocabulary score is small: $\omega^2 = .059$, indicating that about six per cent of the variance in correlations is accounted for by mislead type.

Breakdown of Results According to Part of Speech of the Test Word

It would not be unlikely for the general sequence of stages in word learning to be different for different parts of speech. Among the 26 test words there are fourteen modifiers (adjectives and adverbs) and nine nouns. These two groups have been analyzed separately in the same way as the entire group, and the results are given in Table 5.

Table 5
Correlations of Mislead Types with Total Scores for Nouns and Modifiers

Mislead Type	Correlations with Total (z-trans.)					
	Nouns			Modifiers		
	N	Mean	SD	*N*	Mean	SD
Close	9	.077	.153	13	.035	.129
Context	8	.070	.135	10	-.024	.124
Sound-like	8	-.095	.126	9	-.004	.100
Antonym	8	-.077	.173	14	-.039	.135

The misleads appear to be more strongly differentiated for nouns that for modifiers. However, this could easily be due to sampling error, because the between-groups mean squares for nouns and adjectives are not significantly different from each other: $F(3,3) = 5.26$, where an F value of 9.28 would be required to reach significance at the .05 level. In these very small samples, then, there is no evidence for a difference between nouns and adjectives in how their misleads are ordered.

Effect of Mislead Scoring on Reliability

Given the very small magnitude of the relationship between mislead categories and total vocabulary score, item-option weighting probably would not produce a substantial increase in reliability. If item-option weighting were to be used, the data in Table 3 suggest that an appropriate method would be to give

partial credit for selection of the "close" mislead and no credit for selection of any of the other misleads. It is difficult to say how large the partial credit ought to be, so several values have been tried. The effects on internal-consistency reliability (K-R 20) are shown in Table 6.

Table 6
Reliability with Partial Credit for "Close" Mislead

Points Assigned			
Correct Answer	"Close" Mislead	Other Misleads	Reliability
10	0	0	.591
10	1	0	.595
10	2	0	.597
10	3	0	.597
10	4	0	.593
10	5	0	.585

Note: $N = 326$

Reliability is maximized when the "close" mislead is assigned about one-fourth as much credit as the correct answer. However, as expected, the improvement in reliability is minuscule and would certainly not justify the extra effort required to score both correct answers and close misleads.

DISCUSSION

This study provided evidence that there is a small degree of relationship between mislead categories and vocabulary level. However, the O'Connor hypothesis regarding the way in which mislead categories are ordered was not supported. Rather, the ordering consists principally of a distinction between the "close" mislead and the other three types of misleads. It should not be surprising that the "close" mislead attracts higher-vocabulary examinees because the "close" mislead is, by definition, the mislead most similar in meaning to the test word. The prediction that the antonym might prove to attract examinees even higher in vocabulary than those selecting the "close" mislead was not sup-

ported at all by the data; on average, the antonym was chosen by examinees lowest in vocabulary.

The only implication of the study for our understanding of the process of word learning is the rather obvious idea that the latter stages of learning a word involve making distinctions between the word and other words that are very similar in meaning. If there is a generalizable sequence in the earlier stages of word learning, the present study has not been able to detect it.

Although statistically significant, the difference among mislead categories in their average correlation with vocabulary level is very small in magnitude. In only 13 of 26 items studied does the "close" mislead have the highest correlation with total score (it would be expected to be highest in eight items by chance alone). Also, adding the "close" mislead to the scoring system had virtually no effect on internal-consistency reliability.

The small size of the relationship of mislead types to vocabulary level is also suggested by several multivariate analyses which were performed on a subset of the data (212 cases). Because of technical problems with those analyses, they will only be described briefly. Each examinee was given a set of scores reflecting the number of times he had chosen each of the four mislead types and the correct answer. Both principal-components analysis and factor analysis (using SMCs) were applied to these variables. All methods of analysis showed a large first dimension, corresponding to total-correct score, on which the mislead variables had nearly equal negative loadings; and three smaller dimensions, with approximately equal proportions of variance, which served to segregate the four mislead variables from one another. The similar loadings of the mislead variables on the first dimension would suggest that mislead categories are not related to vocabulary level. There are two major problems in this analysis. First, the five variables are not independent, be-

cause they sum to approximately 26 for each examinee. (Complete dependence is avoided because not all items contain all categories of misleads. Also, dependence can be reduced by using the total-correct score on 150 items rather than on 26 items, but this does not alter the results of the analyses.) Second, the more popular mislead categories will tend to have a greater negative correlation with total-correct score simply because of their greater frequency.

Another multivariate technique, discriminant-function analysis, also was applied. The cases were divided into three groups according to total-correct score on 150 items (high, average, and low). The mislead-category variables were used to discriminate among these groups. Results similar to those for the factor analyses and principal-components analyses were obtained: on the one large discrimant function, the mislead-category variables had negative loadings approximately equal in magnitude. The same technical problems affect this discriminant-function analysis as applied to the previous multivariate analyses.

Although the evidence converges on the conclusion that mislead categories have very little relationship to vocabulary level, it should be pointed out that there are several aspects of this study which would tend to attenuate that relationship. First, the phrase used in the test item may have had a substantial effect on examinees' choice of misleads. Some misleads fit the phrase better than others. Second, not all misleads in a given category may be equally good examples of that category. Third, a substantial number of misleads (27 or 87) attracted less than 10% or more than 90% of examinees who missed the item, and point-biserial correlations for such extremely high or low frequencies are restricted in magnitude. Item-by-item data on response frequencies and correlations are given in Appendices I and II. Because of these limitations in the data, the current study cannot be said to have

shown that there is not a general sequence of stages in word learning, each characterized by a particular kind of confusion. It is possible that a study in which these shortcomings were corrected--perhaps one which obtained free-response definitions from examinees, and which presented the test words alone rather than in a phrase--would reveal such a sequence.

As mentioned previously, the idea that people with a relatively good understanding of a word frequently confuse it with its antonym was probably the most interesting aspect of the O'Connor hypothesis. For some words, such as the pairs port/starboard and stalagmite/stalactite used as examples by O'Connor, the assertion has some intuitive plausibility. Other pairs which might fit with O'Connor's prediction are wax/wane, apogee/perigee, nadir/zenith, yin/yang, and stem/stern. These pairs are either part of a specialized terminology, such as the language or yachting or astronomy, or they refer to opposite ends of a spatial dimension (up-down, left-right, near-far, front-back, enlarging-contracting), or both. It is to be expected that two words which are part of the same technical vocabulary would usually either both be known or both be unknown by a given individual. Also, if two words refer to the two ends of a spatial dimension, and neither word can be used to refer to the dimension itself, it is logical that the two words would tend to be learned at the same time and hence might be confused with each other. Therefore, the words for which confusion with the opposite tends to occur at a late stage of learning appear to be certain types of words which constitute only a small portion of the total lexicon.

For many words, it is obvious why examinees are able to rule out the antonym as the correct meaning even if their knowledge of the test word is relatively poor. The root or prefix of the test word may by itself convey a meaning that is inconsistent with the antonym. Several examples occur in the present study. In the test word DUBITATIVE, the segment "dubi-" suggests the

word "dubious" and contradicts the antonym "confident"; the segment "-spoil-" of the test word DESPOILED suggests the opposite of "prosperous"; and the prefix "mis-" in MISAPPREHENDED is contrary to "grasped-correctly."

In summary, this study has used multiple-choice test data in an attempt to learn something about stages in the process of word learning. The results do not encourage further exploration using this type of data. However, there is a great deal that remains to be discovered about the process of word-learning. Maratsos (1977) described the state of knowledge as follows: "Little in fact is known about the development of lexical meaning...it seems presently implausible to posit developments that are characteristic of the general development of word meaning." Therefore, exploratory research in this area should continue.

APPENDIX I

Percentage of Incorrect Answers Attracted by Each Mislead Category

Mislead Category	Item								
	1	2	3	4	5	6	7	8	9
Close	36.1	45.2	95.1	76.4	32.1	0.0	87.3	33.3	79.1
Opposite	31.3	28.6*	0.0	21.8	42.9	3.5	0.0	15.5	20.9
Context	31.3	--	4.9	0.0	25.0	--	5.1	33.3	0.0
Sound-like	1.2	26.2	--	1.8	--	96.5*	7.6	17.9	--
	10	11	12	13	14	15	16	17	18
Close	44.2	28.2	60.6	42.2	46.9	91.4	32.6	43.0	47.4
Opposite	7.7	7.9	1.8	51.8	0.0	5.5	3.5	34.4	11.6
Context	24.0	--	37.6	3.6	25.0	1.8	63.9	11.8	25.3
Sound-like	24.0	64.0*	--	2.4	28.1	1.2	--	10.8	15.8
	19	20	21	22	23	24	25	26	
Close	55.0	79.6	29.3	24.1	9.1	19.0	51.2	31.3	
Opposite	15.0	5.1	23.3	15.3	51.2*	32.0	31.9	1.4	
Context	3.0	7.1	40.6	5.8	--	49.0	16.9	61.6	
Sound-like	27.0	8.2	6.8	54.7	39.6	--	--	5.7	

* There are two misleads in this category.

APPENDIX II

Point-Biserial Correlation of Each Answer Category with Total Vocabulary Score

Mislead Category	Item 1	2	3	4	5	6	7	8	9
Close	.213	.033	.289	.007	.181	--	.148	-.014	-.035
Antonym	-.326	.117*	--	-.097	.022	.166	--	.024	.035
Context	.064	--	-.289	--	-.220	--	-.175	-.001	--
Sound-like	.175	-.157	--	.278	--	-.166*	-.041	-.003	--
Synonym	.148	.139	.389	.151	.350	.393	.208	.299	.329
	10	11	12	13	14	15	16	17	18
Close	.272	.031	-.045	.387	.044	.075	.076	-.067	-.058
Antonym	-.027	-.082	-.159	-.424	--	-.034	-.248	-.131	-.006
Context	-.166	--	.090	.048	.172	-.053	.020	.359	-.002
Sound-like	-.133	.017*	--	.076	-.214	-.055	--	-.068	.086
Synonym	.271	.399	.373	.285	.289	.051	.281	.369	.392
	19	20	21	22	23	24	25	26	
Close	-.067	.274	-.005	-.120	.060	-.190	.080	-.017	
Antonym	.000	-.139	.130	.177	-.023*	-.047	-.115	-.028	
Context	-.075	-.018	-.034	-.164	--	.193	.036	.055	
Sound-like	.103	-.275	-.144	.052	-.012	--	--	-.067	
Synonym	.242	.038	.202	.004	.142	.136	.346	.140	

* There are two misleads in this category.

Note: For each of the four mislead categories, the correlation is based on the examinees who answered the item incorrectly. For the synonym, the correlation is based on the entire sample of 326 examinees.

REFERENCES

Anglin, J.M. Word, object, and conceptual development. New York: W.W. Norton, 1977.

Bowker, R. Prediction of SAT Verbal scores from General Scale Scores (Statistical Bulletin 1975-30). Boston: Johnson O'Connor Research Foundation, 1975.

Clark, H.H. Space, time, semantics, and the child. In T.E. Moore (Ed.), Cognitive development and the acquisition of language. New York: Academic Press, 1973.

Claudy, J.G. Biserial weights: a new approach to test item option weighting. Applied Psychological Measurement, 1978, 2, 25-30.

Cohen, J. Statistical power analysis for the behavioral sciences (Rev. ed.). New York: Academic Press, 1977.

Donaldson, M. & Wales, R. On the acquisition of some relational terms. In J.R. Hayes (Ed.), Cognition and the development of language. New York: John Wiley & Sons, 1970.

Downey, R.G. Item-option weighting of achievement tests: comparative study of methods. Applied Psychological Measurement, 1979, 3, 453-461.

Kurcaj, S.A. & Maratsos, M.P. On the acquisition of "front," "back," and "side." Child Development, 1975, 47, 202-210.

Maratsos, M.P. Language acquisition: syntactic and semantic development. In B.B. Wolman (Ed.), International encyclopedia of psychiatry, psychology, psychoanalysis and neurology (Vol. 6). New York: Aesculapius, 1977.

Neimark, E.D. Natural language concepts: additional evidence. Child Development, 1974, 45, 508-511.

O'Connor, J. Unsolved business problems. Boston: Johnson O'Connor Research Foundation, 1940.

Rosch, E.H. Human categorization. In N. Warren (Ed.), Advances in cross-cultural psychology, Vol. 1. New York: Academic Press, 1977. (Wickelgren, W.A. Human learning and memory. In Annual review of psychology. Palo Alto: Annual Reviews, Inc., 1981.)

Saltz, E., Soller, E., & Sigel, I.E. The development of natural language concepts. Child Development, 1972, 43, 1191-1202.

Wickelgren, W.A. Human learning and memory. In Annual review of psychology. Palo Alto: Annual Reviews, Inc., 1981.